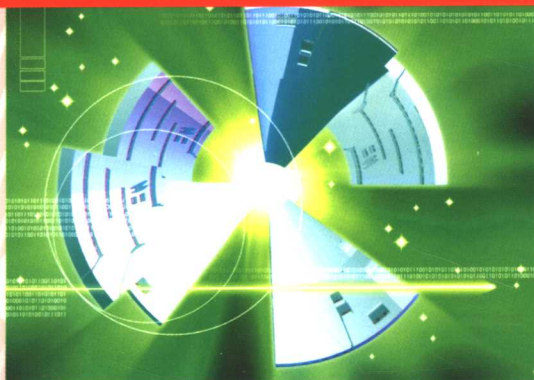


数字时代的图书馆丛书

总编 胡越

数字时代的 图书馆信息资源组织



段明莲 沈正华 编著

 北京图书馆出版社

数字时代的图书馆信息资源组织

段明莲 沈正华 编著

北京图书馆出版社

图书在版编目(CIP)数据

数字时代的图书馆信息资源组织/段明莲,沈正华编著. —北京:北京图书馆出版社,2006.3

ISBN 7-5013-2824-2

(数字时代的图书馆丛书)

I. 数… II. ①段…②沈… III. ①信息资源组织②图书情报工作—研究 IV. G250

中国版本图书馆 CIP 数据核字(2005)第 122291 号

书名 数字时代的图书馆信息资源组织

著者 段明莲 沈正华编著

出版 北京图书馆出版社 (100034 北京西城区文津街7号)

发行 010-66139745 66151313 66175620 66126153

66174391(传真) 66126156(门市部)

E-mail cbs@nlc.gov.cn(投稿) btsfxb@nlc.gov.cn(邮购)

Website www.nlcpress.com

经销 新华书店

印刷 北京华正印刷厂

开本 787×1092 毫米 1/16

印张 17.75

版次 2006年3月第1版 2006年3月第1次印刷

字数 420(千字)

书号 ISBN 7-5013-2824-2/G·635

定价 40.00 元

《数字时代的图书馆》丛书编委名单

胡 越 刘兹恒 朱 强 富 平
陈源蒸 邹 进 宋安莉

总 序

不管学术界如何定义今天这个时代的名词,“数字”和“网络”恐怕是我们所处的这个时代两个最重要的两个主题词。几年前,当这个时代的特征——数字化和网络化扑面而来的时候,曾经在图书馆界引起了一阵争论,甚至恐慌:在数字和网络日益侵占我们的生活的时候,图书馆是否即将步入迟暮之年呢?几年过去了,人们发觉,数字与网络没有取代图书馆的存在,相反,它给这个传统而古老的机构注入了新的活力,使之具有新的生命力。传统的图书馆正在悄悄地进行着深刻的变革。作为图书馆的一名实际工作者,我们时刻关注着、面对着这场变革。很多和我一样工作在图书馆一线的馆员,也在提出这样一些问题:数字时代,图书馆发生了哪些根本的变革?它对图书馆的各项业务工作产生了怎样的影响?图书馆员应该如何改进自己的工作?已有的“数字图书馆”的种种著作,或过多地阐述理论,或过远地描述未来,与图书馆的现实相距甚远。人们迫切需要解决当前问题的务实之作。于是我和陈源蒸老师大胆提出这样一个想法:编写一套专门为图书馆的中、高级专业技术人员培训和继续教育的丛书。想法一经提出,在图书馆圈子里得到大家的首肯,特别是得到人天书店邹进董事长的大力支持,于是有了今天的《数字时代的图书馆》丛书。

在丛书的编写过程中,大家一致认为,该书应该体现如下特点:1. 鲜明的时代特征,强调数字技术对图书馆的影响;2. 突出实践性,力求解决图书馆当前面临的主要问题;3. 观点和材料的新颖性,力求反映最新的研究成果。

本丛书的作者中,既有信息管理专业的专家教授,也有图书馆一线的研究人员,其中不乏德高望重的老教授、当今图情领域里的学科带头人以及在图书馆界崭露头角的年轻学者。他们的加盟使我对这部《数字时代的图书馆》丛书充满了信心。

感谢人天书店邹进、李晓波和北京图书馆出版社的同志,有了他们的远见卓识和细心工作,才有了丛书的顺利出版。

胡 越

2005年9月17日

序 言

信息资源组织是图书馆一项重要的基础工作,是国内外高等院校图书馆与信息管理研究生院开设的专业必修课之一。它包括信息资源的编目、分类、主题标引等内容,在数字化信息资源服务中发挥着日益重要的作用。

随着网络技术和数字出版技术的迅猛发展,图书馆信息资源的组织正面临一些新问题:(1)电子资源迅速增长,图书馆馆藏资源日益丰富多彩。(2)随着数字信息环境本身的发展,信息资源组织的对象、范围、手段和作用也悄然发生变化。图书馆不仅要利用 MARC 元数据或 DC 元数据组织传统的印刷型和声像型信息资源,而且要组织浩如烟海、特征独特的电子资源(如电子图书、电子期刊等)。(3)用户检索需求日益深化,图书馆既要向用户提供整本、整套的信息资源,也要从篇章、图片、镜头、片断等角度向用户深度揭示各种信息资源的内容。(4)数据库数量庞大,数据质量参差不齐,检索平台五花八门,使用对象的目的各异,图书馆正面对如何将各种各样信息资源整合的严峻问题。

因此,我们有必要掌握信息资源组织的基本原理、基本方法和基本手段,更有必要全面地了解网络环境下数字化信息资源组织的新动向和新问题,以便进一步加强信息资源检索工具的功能,不断提高图书馆馆藏资源的查全率、查准率和利用率,不断深化图书馆的服务质量,最终达到共建共享全球信息资源的目的。

本书提纲由段明莲、沈正华、姚星星拟定,经《数字时代的图书馆》丛书编委会集体讨论形成。全书共 11 章,第 1—5 章由段明莲编写;第 6 章由段明莲、张久珍编写;第 7 章由段明莲编写;第 8—9 章由沈正华编写;第 10 章由段明莲、沈正华编写;第 11 章由李冬秀、李燕编写。全书由段明莲统稿。在编写本书的过程中,北京大学信息管理系硕士研究生姚星星为第 11 章提供了大量的参考资料。本书的出版得到了原中宣部陈源蒸老师、国家图书馆富平研究馆员、首都师范大学胡越馆长、北京大学信息管理系刘兹恒教授的指导,得到了北京人天书店有限公司李晓波编辑以及北京图书馆出版社的大力支持,在此,对以上同志一并表示衷心地感谢。

由于时间紧任务重,作者的水平有限,书中难免有疏漏之处,敬请诸位读者批评指正。

段明莲

2005 年盛夏于燕园

目次

序言	(1)
第一章 绪论	(1)
第一节 信息资源概念	(1)
第二节 信息资源的特点	(2)
第三节 信息资源的类型与特征	(2)
第四节 信息资源的组织	(6)
第二章 信息资源编目的基本原理	(11)
第一节 描述性编目的基本知识	(11)
第二节 编目工作的组织模式	(18)
第三节 计算机编目与手工编目的分析比较	(23)
第四节 联机编目的条件与要求	(24)
第五节 联机编目的流程	(26)
第六节 中国文献编目工作的发展趋势	(28)
第三章 MARC 元数据与信息资源组织	(30)
第一节 MARC 的产生与发展	(30)
第二节 CNMARC 与 MARC 21 机读记录的逻辑结构	(32)
第三节 CNMARC 与 MARC 21 的映射关系	(33)
第四节 MARC 元数据的特点	(40)
第五节 MARC 元数据的利弊	(42)
第六节 CNMARC 实施中存在的若干问题	(47)
第四章 DC 元数据与网络资源组织	(52)
第一节 网络资源的特点	(52)
第二节 DC 元数据的产生与发展	(54)
第三节 DC 元数据的设计原则	(57)
第四节 DC 元数据的元素设置	(58)
第五节 DC 元数据的修饰词	(61)
第六节 DC 元数据与 CNMARC 元数据的映射关系	(63)
第七节 DC 元数据与 MARC 元数据的异同	(65)
第五章 电子资源的组织	(69)
第一节 电子资源概述	(69)
第二节 电子资源的组织与揭示	(73)
第三节 数字会议论文的组织	(82)
第四节 电子期刊论文的组织	(89)
第五节 电子学位论文的组织	(94)

第六章 视音频信息资源组织	(103)
第一节 视音频信息资源的特征	(103)
第二节 视音频信息资源组织面临的问题	(105)
第三节 国外视音频信息资源研究状况	(108)
第四节 国外视音频信息元数据研究项目	(114)
第五节 国内视音频信息资源研究状况	(121)
第七章 信息资源的分类标引	(130)
第一节 分类法的种类	(130)
第二节 分类标引的基本原则	(131)
第三节 分类法与主题法的联系与区别	(132)
第四节 《中国图书馆分类法》若干问题	(133)
第八章 信息资源的主题标引	(140)
第一节 主题编目的意义和作用	(140)
第二节 几种主要的主题法	(143)
第三节 主题标引工作	(148)
第四节 分类法与主题法	(150)
第五节 主题法的发展趋势	(157)
第九章 网络环境下主题检索语言的发展趋势	(171)
第一节 FRBR 与主题获取	(171)
第二节 知识组织	(173)
第三节 知识组织系统	(176)
第十章 规范工作	(206)
第一节 规范工作的一般原理	(206)
第二节 规范档的机读格式	(211)
第三节 规范数据库实例介绍与分析	(215)
第四节 规范控制的发展方向	(230)
第十一章 信息资源整合	(234)
第一节 信息资源整合概述	(234)
第二节 信息资源整合的层次与方式	(239)
第三节 信息资源整合的相关技术	(245)
第四节 信息资源整合工具和案例分析	(258)
第五节 信息资源整合的问题及展望	(264)
参考文献	(267)

第一章 绪论

第一节 信息资源概念

“信息资源”(Information resource 或 Information resources)一词源于美国,是随着信息资源管理(Information Resources Management,简称IRM)理论的兴起而产生的术语。国内外关于“信息资源”一词的概念众说纷纭。

早在20世纪70年代末80年代初,美国信息资源管理的创始人——约翰·迭波特认为“信息(或者说信息资源)实质上是数据分析与综合的产物”。^[1]1984年,美国弗吉尼亚州的信息资源管理研究人员里克斯(Betty R. Ricks)和高(Kay F. Gow)在《信息资源管理》一书中指出:“信息资源包括所有与信息的创造、采集、存储、检索、分配、利用、维护和控制有关的系统、程序、人力资源、组织结构、设备、用品和设置。”^[2]¹⁴³美国著名的信息资源管理学家小霍顿(Forest W. Horten, Jr.)认为英语中的“信息资源”一词有单复数之分,其概念也有所不同。“作为单数的信息资源(resource)指信息内容本身,作为复数的信息资源(resources)指各种信息工具包括信息设备、信息用品、信息设施、信息工作者及其信息处理工具,信息财产则指记录在任何媒体上的信息内容(诸如文献、书籍、数据库等)。”^[2]¹⁴⁶而德国信息管理学家斯特洛特曼(K. A. Stroetmann)则认为信息资源由信息内容、信息系统和信息基础结构三部分构成:“①信息内容,它既包括产生于信息服务或从外部信息源获取的信息,也包括与内容活动有关的理论和方法论信息、管理和操作信息、与决策相关的信息,还包括与外部活动有关的交易信息、用户信息和市场信息;②信息系统,其要素包括系统目标、操作人员、信息内容、硬件、内部规则等;③信息基础结构,在此是指一个组织的信息基础结构,它由各种可共享的数据库、计算机硬件设备、数据库管理系统和其他软件、局域网等所构成。”^[2]

我国学者对“信息资源”这一概念的提出比美国晚得多。1991年,中国科学院文献情报中心的博士生导师孟广均教授明确地提出了信息资源的概念,他认为“信息资源包括所有的记录、文件、设施、设备、人员、供给、系统和搜集、存储、处理、传递信息所需的其他机器。”国家信息中心的乌家培先生在《信息资源与信息经济学》一文中对“信息资源提出了两种理解。一种是狭义的理解,即仅指信息内容本身。另一种是广义的理解,指的是除信息内容本身外,还包括与其紧密相连的信息设备、信息人员、信息系统、信息网络等。”我国学者查先进在《论信息资源的含义与经济学特征》一文中也从狭义与广义两个角度阐述了信息资源的概念。查先进在该文中写道“从狭义角度来说,信息资源是指人类社会经济活动中经过加工处理有序化并大量积累后的有用信息的集合,如科学技术信息、政策法规信息、社会发展信息、市场信息、金融信息等,都是信息资源的重要构成要素。从广义角度来说,信息资源是信息和它的生产者及信息技术的集合。也就是说,信息资源由三部分构成:(1)人类社会经济活动中经过加工处理有序化并大量积累后的有用信息的集合;(2)为某种目的而生产有用信息的信息生产者的集合;(3)加工、处理和传递有用信息的信息技术的集合。”

简而言之,信息资源的概念有广义与狭义两种理解。从信息组织对象的角度出发,我国编目界有的学者认为“信息资源是反映客观事物的各种信息的总称,也是各种信息的集合体。”从信息资源管理角度出发,有的学者的观点则倾向于“信息资源包括所有与信息的创造、采集、存储、检索、分配、利用、维护和控制有关的系统、程序、人力资源、组织结构、设备、用品和设置。”本书采纳了信息资源的狭义概念,只涉及记录在各种媒体上的信息资源(包括文字、图像、声音、数据等)的组织问题。

第二节 信息资源的特点

当今世界已步入信息时代。信息化社会的最重要特征是“信息资源”和自然资源一样,是社会发展的不可缺少的资源,是一种宝贵的、富有价值的资源。信息资源与自然资源相比,具有下列特点:

1. 信息资源是一种动态资源,产生于自然界和人类社会的实践活动之中,是随着时间的变化而变化的。人类的社会实践活动是一个永不停歇的运动过程,信息也总处在不断产生、积累的过程之中,并呈现出不断丰富、不断增长的趋势。信息资源是在人类参与的过程中不断产生、不断完善的。因而,信息资源的再生性很强,重复使用不会损耗或减少其价值。

2. 信息资源是经过人类开发与组织的资源,并非任何信息都能成为资源,只有经过人类开发与组织的信息才是资源。因此,信息资源是经过人类开发与组织的信息的集合,是通过人类的参与而获取的信息,而开发与组织正是信息资源可利用性的表征。

3. 信息资源是一种可供利用和共享的资源。信息资源中所包含的各种信息既可以为决策服务,又可以为经济建设、科学研究和人类社会生活服务。在可利用性方面,信息资源在计算机网络中不仅可以不受国界、地区、政治、经济、文化和语言的限制进行传输,而且可供多位用户共同使用,是一种可以共享的资源。

4. 信息资源是一种具有依附性的资源。其依附性表现在两方面:一个是信息资源的传递需要专门的传播手段,另一个是文献信息资源与信息载体具有不可分性。

第三节 信息资源的类型与特征

信息资源是信息组织的对象,是满足人类生存需求资源之一。信息资源的类型繁多,不同划分标准可划分出不同类型的信息资源。按照出版技术划分,信息资源可分为非电子资源和电子资源;按照内容划分,信息资源可分为视频信息资源、音频信息资源以及视音频信息资源三类;按照表现形式划分,又可以分为静态信息资源和动态信息资源。

一、非电子资源

根据《中国文献编目规则》第2版,非电子资源包括图书、学位论文、科技报告、标准文献、古籍、拓片、测绘制图资料、乐谱、录音资料、影像资料、静画资料、连续性资源、缩微文献、手稿。本节着重介绍以下几类信息资源。

1. 图书

图书(books)是一种以印刷形式刊行的,不少于49页的,有正文或插图的非连续出版的出版

物。通常由书衣、封面、书脊、文前栏目、扉页、半书名页、附加书名页、书名页、版权页、献词、序言或前言、目次、绪论或导论、正文、图表、附录、参考文献、索引这几部分组成。印刷型图书将信息记载在轻便耐久的纸张上,阅读方便,其目的是在人与人之间传递信息、传授知识、记录历史、交流沟通思想感情。它不受时间和空间的限制,行使着宣告、阐述、保存与传播知识的职能。

正式出版的图书载有的国际标准书号(International Standard Book Number,简称ISBN)由冠有“ISBN”字符的10位数字组成。ISBN分为组号(Group identifier)、出版者号(Publisher identifier)、书序号(Title identifier)以及校验号(Check digit)四部分,各个部分之间置连接符或空格。例:ISBN 0-87287-220-3。此例中“0”为组号,“87287”为出版者号,“220”为书序号,“3”为校验号。

1986年,我国参照ISO 2108—1978《文献工作——国际标准书号》制定了中华人民共和国国家标准GB 5795-86《中国标准书号》。中国标准书号与国际标准书号的区别在于在ISBN后附了《中国图书馆分类法》基本大类的分类号和种次号。

例:ISBN 7-5013-1136-6/G·307

“ISBN 7-5013-1136-6”是国际标准书号,其中“7”为组号,“5013”为出版者号,“1136”为书序号,“6”为校验号,“G”为分类号,表示《中国图书馆分类法》的“文化、教育、体育”基本大类,“307”为种次号。

2. 连续性资源

连续性资源(continuing resources)是一种预先未明确终止时间、连续出版发布的资源和不断更新的一体化资源,包括连续出版物和不断更新的活页本、网站等。在此说明一点,网站是一种网络资源,属于电子资源的范畴,而不属于非电子资源的范畴。

连续出版物(serials)是一种载有编号或年月顺序号,计划无限期地连续出版发行的印刷型的或非印刷型的出版物。它包括期刊、报纸、年刊、学会报告丛刊与会刊、会议录、科学进展、年鉴以及专著丛刊等,但不包括在一个预定期限内以连续分册形式发行的丛书和多卷书。连续出版物具有几个显而易见的特征。其一,出版的连续性。连续出版物的主办者、编辑者力求用一个稳定的总题名使连续出版物无限地连续不断地出版。其二,具有期标识系统、出版周期和国际标准连续出版物号(International Standard Serial Number,简称ISSN)。其三,版式设计的稳定性。从历史的角度分析问题,同一连续出版物的各卷/期的题名、责任者、出版者、开本、期标识、封面设计、篇幅、内容范围等方面具有一致性。其四,内容新颖、时效性强。有的连续出版物不仅内容新颖,而且综述了一定时期内某一科学领域的发展和成就。有的连续出版物甚至是连续出版的检索工具,内容及时,信息密集,叙述简明,编排得当,便于读者检索信息资源。

3. 测绘制图资料

测绘制图资料(cartographic materials)又称舆图资料,是指以任何比例全面地或局部地绘制、显示地球或其他天体的资料。它包括平面地图、立体地图、航空图、航海图、星体图、街区图、球仪、航空照片、卫星照片、地图集、鸟瞰图等等。人们最常见的是平面地图与地图集。

测绘制图资料是一种特殊类型的信息资源,具有特殊的数学法则,需使用地图学语言描述此类文献的特性,通常有比例尺、投影法、经纬度、色彩等特征。地面上各种地理事物,不可能按它们的实际大小表示在图纸上,而是需要缩小一定的倍数来进行描绘。这种缩小的比例关系就是地图的比例尺。地图比例尺一般分为大比例尺,中比例尺和小比例尺三种。众所周知,地球椭球面是一个不可展的曲面,将地球椭球面上的点投影到平面上来的方法称为投影法。

地图投影按投影变形性质划分,通常将投影法分为等角投影、等积投影和任意投影;按正常位置投影经纬网图形划分,可分为方位投影、圆柱投影、圆锥投影、伪圆柱投影等。测绘制图资料通常用经度与纬度表示某地的方位。此外,测绘制图资料用各种色彩表示不同的地貌类型和区划,运用色彩一方面是因为它具有美学上的优点,更重要的是因为它有增加清晰性的效果。

4. 录音资料

录音资料(sound recordings)又称录音制品,是指“用机械或电子手段,声音振动转录于录音材料(磁带、磁盘、唱片、胶片等等)之上形成的文献”,^{[3]539}包括唱片、录音带(如,循环录音带、盒式录音带、开盘录音带)、声道胶片、激光唱盘以及MP3。通常由容器、标签、正文以及附件组成。其特点在于:第一,提供音频信息;第二,载有播放时间、转速、声道数、ISBN号以及ISRC号(如:ISRC CN-A09-98-304-00);第三,借助录音机、唱机等专用的设备收听视频信息。

5. 影像资料

影像资料(motion pictures and video recordings)包括电影制品和录像制品。电影制品又称影片,是指载有连续图像通过快速投影给视觉造成动感的透明胶片,包括盒式循环电影片、盒式电影片、环式电影片以及开盘电影片。录像制品又称录像资料,是指将视音频信息以模拟信号或数字信号的方式记录在某种载体上,通常是活动图像有伴音,并借助特定设备放映的制品。它包括循环录像带、盒式录像带、开盘录像带、数码流带以及视盘(如:激光数码视盘、高密度激光视盘)等等。其特点在于:第一,提供视音频信息。第二,借助专用设备将信息直接传入用户的耳膜和视网膜,从而达到“阅读”的目的。这种能使人观其形、听其声,给人以动态直观的感觉的信息资源在帮助人们学习语言知识、观察科技现象、传播科学技术、娱乐消遣等方面,发挥着独特的作用。第三,通常载有时长,节目入点、出点,音响,色彩,投影速度等。第四,存储和保管要求高。不仅须防止磁化和老化,而且注意温度和湿度,否则就会发脆、变形、断裂。

6. 缩微文献

缩微文献(Microforms)又称缩微制品。它是原件的复制品,是窗孔卡(aperture card)、缩微胶片(microfiche)、缩微胶卷(microfilm)以及缩微卡片(microopaque)的总称。通常由容器、题名帧、正文、附件等部分构成。其特点在于:第一,通常载有缩率、色彩,并分正片与负片。第二,体积小,重量轻,信息量大。第三,阅读条件受限,必须借助缩微设备方可阅读。

7. 静画资料

静画资料(Graphic Materials),又称图形制品、图示资料,是指不透明的、以二维方式表现的画像,或者是一种通过光学设备观看或投影的、以二维方式表现的静态图形。它包括美术原作(art original)、美术印刷品(art print)、美术复制品(art reproduction)、图表(chart)、活动挂图(flip chart)、照片(photograph)、图片(picture)、幻灯片(slide)、工程图(technical drawing)、投影片(transparency)以及挂图(wall chart)等等。图形制品的特点在于:第一,刊载的内容图文并茂,所反映的图形形象直观、一目了然。不仅可以直观揭示所描述的对象的形象特征,而且可以表达作者的创作思想。第二,图形制品通常是以二维方式与静态的表现手法揭示图像内容。第三,在物质外形特征方面,图形制品的通常载有创作者、绘制者、创作日期或绘制日期等。第四,图形制品的介质既有纸张,也有感光材料。

二、电子资源

1. 电子资源的类型

电子资源(Electronic resources)是一种数字信息资源。由于科学技术的发展,特别是进入20世纪80年代后,计算机技术和网络技术飞速发展,致使电子出版物异军突起并呈现急速增长的势头。电子出版物以收藏占用空间小、信息存储量大、信息检索便捷等特点已为人们重视与利用。

“电子资源”一词的使用有个演变过程。《英美编目条例》第2版1988年修订本最初将电子出版物归入计算机文档(Computer files)。1990年国际图书馆协会联合会主持制定了《国际标准书目著录(计算机文档)》(International Standard Bibliographic Description for Computer File,简称ISBD(CF))。ISBD(CF)认为计算机文档是“经过编码处理可用计算机管理的资料(数据文献或程序)”。针对近年来电子信息技术迅速发展,电子资源种类及范围不断增加,1994年经全世界的评论,《国际标准书目著录(计算机文档)》第2版扩充了内容,并取名为《国际标准书目著录(电子资源)》(International Standard Bibliographic Description for Electronic Resource,简称ISBD(ER))。随着电子资源的种类和范围不断发展,《国际标准书目著录(电子资源)》和《英美编目条例》第2版2002年修订本都用“电子资源”一词取代了“计算机文档”,因为“电子资源”一词比“计算机文档”能更为全面地包容此类资源。

关于电子资源的类型,由于划分的标准不同,可得出不同分类结果。

按照发布形式划分,电子资源可分为电子图书(e-book)、电子杂志(e-journal)、电子报纸、电子声像、数据库等等。

按照存取方式划分,电子资源可分为远程访问的电子资源和近程存取的电子资源。远程访问的电子资源,又称网络访问的电子资源,是指不必凭借物理载体,只需通过启动某一特定的计算机就可调用存储在硬盘上的指定文档,或者利用计算机网络、远程通信等方式获得计算机文件。近程存取的电子资源,又称本地访问的电子资源,是指通过计算机设备或外围设备的物理载体获取的信息资源,其物理载体为光盘、磁盘、磁带、盒式磁带等。例如,计算机软磁盘、计算机磁带、计算机光盘等。

按照存储的信息特征划分,电子资源可分为文本信息、图形信息、图像信息、音频信息、视频信息等等。

按照存储内容划分,《国际标准书目著录(电子资源)》将电子资源划分为电子资料、电子程序以及电子资料及程序三大类。电子资料(Electronic data)包括:电子数据资料,如,电子人口普查资料、电子调查资料;电子文件资料,如,电子书目数据库、电子期刊、电子会讯、电子文献;电子影像资料;电子表述资料,如,电子的地图资料、电子的音乐资料;电子声音资料;电子字型资料。电子程序(Electronic program(s))包括:电子公用程序;电子应用程序,例如,计算机辅助绘图程序、电子试算程序、电子文字编辑程序、电子桌上排版出版程序、计算机游戏;电子系统程序,如,电子操作系统程序、电子程序语言、电子检索程序。电子资料及程序(Electronic data and program(s))包括电子交互式多媒体和电子在线服务。

2. 电子资源的特点

电子资源与非电子资源相比,尽管它们的主要功能都是传递信息,但在许多方面有其独具一格的特点,主要表现在以下几个方面:

(1) 信息载体别具一格

非电子资源的物理载体主要是纸张、感光材料或磁性材料,而电子资源的载体则是磁盘、磁带、光盘、硬盘,甚至可以通过Internet网传播。刊载在纸张上的印刷型文献的内容具有难

以修改更新的弱点,而电子资源具有可重复读写的特性,易于修改、更新、重用,因而,存储在可读写的磁盘、光盘、硬盘上的信息容易被篡改,可靠性差。

(2) 数据存储量大,信息丰富多彩

与印刷型文献相比,电子资源具有数据存储量大的特点。现以近程存储的电子出版物为例,普通 CD 的存储容量为 600MB 或 750MB。DVD 视盘存储量比 CD-ROM 大得多,单面单层 DVD 视盘的数据容量为 4.7GB,是 CD 容量的 9 倍。双面双层的 DVD 视盘可达 17GB 的容量,这使高清晰画面和高品质音响的存储与处理都显得游刃有余。一张光盘可以容纳几十亿个字符,一张 SDVD 可以存储 3000 首卡拉 OK,甚至可以容纳 10 集电视连续剧。由此可见,电子资源不仅数据存储海量,能记录文字和静态图像,而且可以利用多媒体技术将文字、图形、声音、动画、图像等集于一体,向用户提供丰富多彩、生动逼真的信息,使用户能够听其声、观其形,犹如亲临其境。

(3) 以超文本的形式组织信息

图书、报刊、电视、广播等是以线性形式组织信息,用户用这些信息资源时,只能依据所提供的的使用方式,顺序获取所需信息。对于印刷型图书,用户查找信息时,须通过不断地翻页,不断浏览全书获取所需信息。而电子资源,尤其是网络信息源不仅可以用线性形式组织信息,而且可以超文本的形式组织信息,所有的信息通过超文本链接形成网状结构,用户可以从一节点链接到其他节点。超文本链接是电子资源区别于传统信息源的一大特点。

(4) 检索系统完备,检索路径全面

传统的印刷型文献通常只能向用户提供题名索引、著者索引、主题索引以及分子式索引。而电子资源不仅可以提供题名、著者、主题词、关键词、分类号等多种单项检索,而且可以运用布尔逻辑检索技术进行多项组配检索,甚至可以提供截词检索、全文检索以及图像检索,向用户提供信息资源的精确检索或模糊检索,其信息检索的广度和深度可以随不同的检索目的而改变。因此,电子资源在文献的查全率、查准率以及检索速度上都具有得天独厚的优势。

(5) 携带方便,阅读条件受限

印刷型文献体积大,但阅读条件不受时空限制,可以在任何时间、任何地点进行。而电子资源体积小,携带方便,但电子资源的利用受获取条件的制约,必须借助特定的设备,远程存取的网络资源还需计算机网络的支持才能实现。

第四节 信息资源的组织

一、信息组织的内容

信息组织是根据用户检索信息资源的需要,以文本、图像等各种信息资源为对象,通过对物质外形特征和内容特征的分析、选择、标引、记录,使其成为有序化集合的活动。信息组织工作包括信息资源的描述、主题标引、分类标引、规范文档的建设、检索工具或检索系统的建立。信息资源描述亦称描述性编目(Description cataloguing),是指以编目记录的形式对信息资源的物质外形特征进行分析、选择和记录,并根据信息资源的外在特征赋予其检索标识的过程。信息资源的主题标引与分类,亦称主题编目(Subject cataloguing),是指根据信息资源的主

题内容和其他特征,赋予其相应的分类号或主题标识的过程。

信息组织的根本目的在于:将杂乱无章的信息组织成一个有序化的集合体,满足用户检索信息资源的各种需求,从而达到获取信息资源的目的。图书馆对信息资源进行组织需要借助文献著录规则(例:《中国文献编目规则》、《西文文献著录条例》等)、主题法(例:《汉语主题词表》等)、图书分类法(例:《中国科学院图书分类法》、《中国图书馆分类法》等),计算机编目还需借助图书馆自动化集成系统和机读目录格式(例:《中国机读目录格式使用手册》、USMARC Format 等)。

二、信息组织的工具

信息组织需要借助文献编目规则、主题法、图书分类法、机读目录格式以及图书馆编目系统。本节着重介绍 GB 3792 系列国家标准、《中国文献编目规则》、AACR2 以及《西文文献著录条例》这几个信息组织工具。

1. GB 3792 系列国家标准

1979 年 12 月,经国家标准总局批准,正式成立了全国文献工作标准化技术委员会(现更为全国信息与文献标准化技术委员会)。该委员会下属的第六分委员会在 1983—1987 年期间主持制定了用于文献信息资源描述的国家标准有:

- 《文献著录总则》(GB 3792.1—83)1983 年 7 月 2 日发布,1984 年 4 月 1 日实施;
- 《普通图书著录规则》(GB 3792.2—85)1985 年 1 月 31 日发布,1985 年 10 月 1 日实施;
- 《连续出版物著录规则》(GB 3792.3—85)1985 年 2 月 12 日发布,1985 年 10 月 1 日实施;
- 《非书资料著录规则》(GB 3792.4—85)1985 年 2 月 12 日发布,1985 年 10 月 1 日实施;
- 《档案著录规则》(GB 3792.5—85)1985 年 5 月 10 日发布,1986 年 1 月 1 日实施;
- 《地图资料著录规则》(GB 3792.6—86)1986 年 6 月 19 日发布,1987 年 1 月 1 日实施;
- 《古籍著录规则》(GB 3792.7—87)1987 年 1 月 3 日发布,1987 年 10 月 1 日实施。

1999 年,全国信息与文献标准化技术委员会第六分委员会重新修订了《文献著录总则》、《普通图书著录规则》、《连续出版物著录规则》以及《非书资料著录规则》。2000 年这四项国家标准报批修订稿报送全国信息与文献标准化技术委员会。此后,全国信息与文献标准化技术委员会第六分委员会还根据 ISBD(ER)制定了《电子资源著录规则》。这一切都促进了我国文献著录标准化发展进程,为文献信息资源的共建共享奠定了坚实基础。

2. 《中国文献编目规则》

GB 3792 文献著录标准问世后,中国图书馆学会根据此套国家标准编撰了《中国文献编目规则》。该编目规则于 1995 年由广东人民出版社出版,其内容包括著录法和标目法两大部分。

2005 年,国家图书馆《中国文献编目规则》修订组根据 ISBD 和 AACR2 的最新版本,并结合中国文献编目工作的实际需要,对 1995 年出版的《中国文献编目规则》的规定信息源、版本信息选取、特殊文献著录对象的确立、分析著录的不同类型等规则进行了修订,同时对电子资源、连续性资源、测绘制图资料等章节做了补充与修改,特别是标目法部分,在内容、章节、结构上做了较大调整,并增加大量样例,使其内容更加充实。

3. 《英美编目条例》第 2 版

早在 1978 年,美国图书馆协会等单位分别出版了《英美编目条例》第 2 版(Anglo-American Cataloguing Rules 2nd ed.,简称 AACR2);1988 年出版了《英美编目条例》第 2 版 1988 年修订本;

2002年出版了《英美编目条例》第2版2002年修订本;预计在2008年将出版AACR3。英美编目条例合作筹划指导委员会(The Joint Steering Committee)在AACR3草案中,简明扼要地阐述了范围、结构、术语以及著录规则描述功能的目的与原则。在范围、结构、术语等方面,奉行总则一般化、补充规则专门化、术语与FRBR保持一致、减少冗余等原则,其目的在于:第一,确保编目规则的范围、结构以及术语等的广泛性、一致性、合理性、兼容性、适应性;第二,使用简易与效率;第三,能够引见印刷型或数字格式。在描述功能方面,奉行区分各类资源、满足用户需求、加强资源间的关联等原则,其目的在于:提高用户需求的相应性、成本效率、格式的独立性。根据AACR3修订原则和目的,它将进行重大修改。例如,将题名《英美编目条例》改为《资源描述与检索》(Resource Description and Access,简称RDA)。又如,编目条例第一部分的结构将会有重大变化。著录部分将改变总则和分则的格局,著录部分的结构将包括导论、总则、适用于特定内容类型的补充规则以及适用于特定媒体类型的补充规则这四部分内容。

4.《西文文献著录条例》

AACR2的产生与发展对我国《西文文献著录条例》的制定影响颇大。1983年8月,为了进一步探讨西文图书著录标准化问题,全国高等学校图书馆工作委员会与全国信息与文献标准化技术委员会第六分委员会在北京召开了“西文图书编目标准化与自动化研讨会”。与会者由编目工作者、西编教员以及图书馆自动化系统研制人员构成,分别来自各种类型图书馆和高等院校教学部门。会议充分地研讨了ISBD的推广、AACR2的应用以及文献编目的自动化问题。会后,根据研讨会的提议成立了中国图书馆学会《西文文献著录条例》编辑组。1985年,中国图书馆学会本着等效采用ISBD、“基本采用,个别改动”AACR2的精神,结合我国西编工作的实际需要制定,出版了《西文文献著录条例》。经过十多年的实践,中国图书馆学会《西文文献著录条例》修订组修订了《西文文献著录条例》。科学技术文献出版社于2003年7月出版了《西文文献著录条例》修订扩大版。该条例的宗旨是“结合我国实际需要,采用了AACR2,1998修订本和AACR2R—93和ISBD国际标准原则,试图既能满足手工记录的需求,又能照顾到自动化的发展,达到书目记录共享的目的”。^{[3]:2}

三、信息组织的原则

1. 信息资源描述的标准化

信息资源描述的准确性是指受编信息资源的著录结果完全符合自身的物质外形特征与其内容的实际情况。它是信息资源编目科学性的具体表现,是提高信息资源检索和书目数据库质量的保证。

机读书目记录和传统的款目是通过每一种信息资源的标引、著录,而获得每种信息资源的内容和物质外形特征的记载。借助这些特征来确认、了解、分析比较信息资源,从而达到获取信息资源的目的。为了充分地揭示信息资源,方便用户检索利用,就必须全面、客观、准确地揭示信息资源。这就意味着编目工作者著录信息资源时,需依据文献著录规则揭示信息资源的物质外形特征,反映信息资源的内容。

1969年,为了统一各国的书目著录,在丹麦的哥本哈根召开了国际编目专家会议。会议一致认为应当统一各国书目著录信息,建立国际信息交换体系。此后,国际图书馆协会联合会(International Federation of Library Associations and Institutions,简称IFLA)的编目委员会制定了一套供各类型信息资源使用的《国际标准书目著录》(International Standard Bibliographic

Description, 简称 ISBD)。到目前为止,已出版的 ISBD 有:《国际标准书目著录(总则)》、《国际标准书目著录(专著)》、《国际标准书目著录(连续出版物)》、《国际标准书目著录(舆图资料)》、《国际标准书目著录(非书资料)》、《国际标准书目著录(乐谱)》、《国际标准书目著录(古籍)》、《国际标准书目著录(分析著录)》以及《国际标准书目著录(电子资源)》。1983 年后,全国信息与文献标准化技术委员会第六分委员会参照 ISBD 相继制定了一系列信息资源著录方面的国家标准。2000—2001 年,第六分委员会分别修订了 GB 3792.1《文献著录总则》、GB 3792.2《普通图书著录规则》、GB 3792.3《连续出版物著录规则》、GB 3792.4《非书资料著录规则》这几个国家标准,并依据 ISBD(ER)制定了《电子资源著录规则》。

全面准确地著录信息资源,有利于图书馆目录如实地揭示信息资源,有利于用户检索、识别、选择信息资源,有利于确保书目记录的准确性,有利于确保书目数据库的质量,提高信息检索的查准率。

2. 检索点的规范化

检索点(Access point)是指用于查找和识别一条书目记录的名称、术语、代码等,是用户进行信息检索的入口,具有检索和识别书目记录的功能。由于社会、政治、历史、婚姻等原因致使个人著者的姓名、机关团体的名称乃至文献题名等发生一些变化,生成各种各样的形式。因而,检索点的选取必须对准信息资源的基本特征和用户查找信息资源的口径。

为了加强检索点的规范化,确保检索点的一致性与惟一性,美国早在 1977 年开始实施“名称规范合作计划”,开展了书目记录的规范化工作。规范工作(Authority work)是建立、维护、使用和评估规范记录的过程,是确保机读书目记录检索点(如,个人著者姓名、机关团体名称、题名及主题等)规范化的具体措施,是创建高质量书目数据库的一项重要工作。我国于 20 世纪 90 年代开始重视此项工作。中国国家图书馆(National Library of China)、北京大学图书馆(Peking University Library)等单位先后建立了规范文档。倘若利用中国国家图书馆的书目查询系统,从“孙中山”著者姓名的角度检索文献,一次就可检出 37 条机读记录,既可以检到“孙中山”著的《孙中山选集》,也可以检到以“孙文”这一姓名著的《建国方略》。检索点的规范工作与文献著录标准化工作一样需要加强馆际间合作与协调。只有这样,才能在书目数据库中集中同一责任者的不同信息资源,汇集同一信息资源的不同版本、译本、文本或载体,确保信息检索的查全率和查准率。

3. 主题标引的专指性

为了确保主题标引的专指性,信息资源一方面要根据其内容进行标引,另一方面要用最专指的主题词标引。“当主题词表收有多个相关主题词时,应选词表中与信息资源的主题概念相对应的最专指的主题词进行标引,而不得以该词的上位词或下位词进行标引,以免出现标引过宽或过窄的误差。”^[4]专指度是主题标引的一项查词规则,当词表中没有相应专指主题词时,可依次选用组配标引、上位叙词标引、靠词标引以及增词标引等方法揭示信息资源的内容,加强主题标引专指性的根本目的在于提高信息资源的查准率。

4. 信息资源的揭示程度

信息资源的揭示程度包括受编信息资源的揭示层次和主题标引的深度。信息资源的揭示层次分为整体揭示和分析揭示两类。整体揭示是以种为单位描述信息资源的特征和内容。分析揭示是以篇章、图像、视音频信息资源的片断或镜头为单位描述析出信息的特征和内容,是深层次揭示信息资源的一种方法。主题标引可分为浅标引和深度标引,浅标引采用概括分析